Centers for Disease Control and Prevention
Epidemiology Program Office
Case Studies in Applied Epidemiology
No. 811-703

Oral Contraceptive Use and Ovarian Cancer

Instructor's Guide

Learning Objectives
After completing this case study, the participant should be able to:

☐ Outline the sequence of an epidemiologic analysis;

☐ Discuss the biases of particular concern in case-control studies and ways to minimize their influence;

☐ Describe why and when to use crude and adjusted odds ratios and 95% confidence intervals, and how to interpret them; and

☐ Define and recognize effect modification and confounding.

This case study was developed by Richard Dicker and Peter Layde in 1981. Current version updated by Richard Dicker with input from the EIS Summer Course instructors.

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service

PART I

In 1980, ovarian cancer ranked as the fourth leading cause of cancer mortality among women in the United States. An estimated 18,000 new cases and more than 11,000 attributable deaths occurred among American women that year.

Several studies had noted an increased risk of ovarian cancer among women of low parity, suggesting that pregnancy exerts a protective effect. By preventing pregnancy, oral contraceptives (OCs) might be expected to increase the risk of ovarian cancer. On the other hand, by simulating pregnancy through suppression of pituitary gonadotropin release and inhibition of ovulation, OCs might be expected to protect against the subsequent development of ovarian cancer. Because by 1980 OCs had been used by more than 40 million women in the United States, the public health impact of an association in either direction could be substantial.

To study the relationship between oral contraceptive use and ovarian cancer (as well as breast and endometrial cancer), CDC initiated a case-control study – the Cancer and Steroid Hormone (CASH) Study in 1980. Case-patients were enrolled through eight regional cancer registries participating in the Surveillance, Epidemiology, and End Results (SEER) program of the National Cancer Institute.

Question 1: Which investigations need to be reviewed by an institutional review board? Does this investigation need to be reviewed?

Answer 1
Basic concept is whether the purpose is to benefit the persons or community in the study or to generate "generalizable knowledge." The purpose of this investigation is clearly to learn more about the association between OCs and ovarian cancer, i.e., generalizable knowledge, rather than to directly benefit the participants. Therefore, this study must be reviewed by an IRB.

As the investigators planned this study, they discussed a variety of methods to minimize potential biases.

Question 2: What types of bias are of particular concern in this case-control study? What steps might you take to minimize these potential biases?

Answer 2
In case-control studies, recall bias and the various types of selection bias are of greatest concern. In multicenter studies, selection, interviewer, and data-collection biases are of added concern.

Instructor's note: Do not belabor this question! The following information is provided for instructors. You do not need to cover all this material with the students.

Bias is a systematic error in the design or conduct of a study that results in a distorted estimate of an exposure's effect on the risk of disease (or other health problem).

Systematic error (bias) is related to the design or conduct of a study. In contrast, random error is related to sample size.

In general, biases can be categorized as follows:

I. **Selection bias**: how the subjects get into the study. Types include:
   - Detection (also known as surveillance bias)
   - Referral (hospitalization, Berksonian)
   - Nonresponse or self-selection
   - Inappropriate comparison group
2. **Information bias**: after the subjects have been selected, how exposure information is collected
   - Recall
   - Interviewer/abstractor
   - Data collection
   - Prevarication (lying)
3. **Confounding**

Methods to reduce bias

I. Selection bias
   a. Case-patients — Use a "clean," confirmable, consistent definition of a case, using objective criteria if possible. Try to enroll all case-patients in a defined time and region.
   b. Controls — Controls should come from the same population as the case-patients, so that they represent the exposure prevalence in the population from which the cases came. A key in selecting a control is that if the person got the disease, the person would have been identified as a case-patient. (Most epidemiologists consider that controls must be at risk for the disease to be eligible; however, this is a current topic of debate.) Select controls at random if possible.

2. Information bias
   a. Recall — Use memory aids. If possible, validate exposure data with independent sources.
   b. Interviewer, data collection — Data should be collected from case-patients and controls in the same manner, using a standard data-collection form to abstract records, conduct interviews, etc. Interviewers (or abstractors) should be trained so that they conduct interviews uniformly (minimizing both intra- and inter-interviewer variability). When possible, interviewers should not know case-control status of person being interviewed.

As the investigators began to consider what data to collect with their questionnaire, they began to lay out the analyses they wanted to conduct. They did so by sketching out "table shells" -- frequency distributions and two-way tables that contain no data but otherwise include appropriate titles, labels, measures and statistics to be calculated. The tables followed a logical sequence from the simple (descriptive epidemiology) to the more complex (analytic epidemiology) that is often used when results are presented in a manuscript or oral presentation.

Question 3: List, in logical sequential order, the table shells you might use to analyze or present the CASH study data.

Answer 3

INSTRUCTOR'S NOTE: Break up into groups of about 4 for this question.

The point of this question is that analysis should start with a plan. It should proceed from descriptive ("getting to know your data") to analytic, from simple to complex. The sequence described below is appropriate for any type of epidemiologic study.

Prelim Table (sometimes seen, especially in clinical trials): enrollment statistics (number eligible, number enrolled, not enrolled by reason, e.g., died, refused, etc,

Table 1:	Clinical, e.g., frequency distribution by types of ovarian cancer.
Table 2:	Descriptive ("person, place, ±time") -- who is in the study, e.g., frequency distributions by case-control status for demographics (age, race) and ovarian cancer risk factors (marital status, parity, infertility).
Table 3:	Primary 2-by-2 table: OC use (ever/never) by case-control status.
Table 4+:	Stratification of Table 3 by covariates such as age, race, and parity to assess confounding, effect modification.
Table 5+:	Refinements of Table 3, e.g., dose-response, latency, etc.
Table 6+:	Specific subgroup analyses.

PART II

The study design included several features to minimize selection and information bias. Ascertainment bias of disease status – a type of selection bias – was minimized by attempting to enroll as cases all women ages 20-54 years with newly diagnosed, histologically confirmed, primary ovarian cancer who resided in one of the eight geographic areas covered by the cancer registries. Controls were women ages 20-54 years selected randomly using telephone numbers from the same geographic areas. Because 93% of U.S. households had telephones, virtually all women residing in the same areas as the cases were eligible to be controls. (Interestingly, all the women enrolled with ovarian cancer had telephones.)

To minimize interviewer bias, CDC investigators conducted group sessions to train interviewers in the administration of the pretested standard questionnaire. The same interviewers and questionnaires were used for both cases and controls. Neither cases nor controls were told of the specific a priori hypotheses to be tested by the study. Recall bias of oral contraceptive exposure was minimized by showing participants a book with photographs of all OC preparations ever marketed in the United States and by using a calendar to relate contraceptive and reproductive histories to other life events.

The primary purpose of the CASH study was to measure and test the association between OC use and three types of reproductive cancer – breast cancer, endometrial cancer, and ovarian cancer. Enrollment of subjects into the study began in December 1980. During the first 10 months of the study, 179 women with ovarian cancer were enrolled, as well as larger numbers of women with endometrial or breast cancer. During the same period, 1,872 controls were enrolled to equal the number of subjects with breast cancer. The same control group was used for the ovarian cancer analysis; however, the investigators excluded 226 women with no ovaries at the time of interview and four controls whose OC use was unknown, leaving 1,642 women to serve as controls. The distribution of exposure to OCs among cases and controls is shown in Table 1.

Table 1. Ever-use of oral contraceptives among ovarian cancer cases and controls, Cancer and Steroid Hormone Study, 1980-1981

		CASE-CONTROL STATUS		
		Case	Control	Total
USE OF OCs	Ever	a = 93	b = 959	H_1 = 1052
	Never	c = 86	d = 683	H_0 = 769
	Total	V_1 = 179	V_0 = 1642	T = 1821

Question 4: From these data, can you calculate the risk of ovarian cancer among oral contraceptive users? Why or why not?

Answer 4
No. The risk of disease is defined as the proportion of an initially disease-free population who develop the disease during a specified period of time. *Attack rate* is a synonym for risk commonly used in acute outbreak settings.

The row totals in this two-by-two table represent the number of exposed and unexposed subjects in the study, but not the size of the disease-free population from which the cases arose. These row totals are artificial and arbitrary, because the investigators can select the numbers of controls per case. The actual size of the exposed and unexposed populations that gave rise to the cases is rarely known in case-control studies.

Question 5: Describe the rationale behind using the odds ratio as an estimate of the risk ratio. When is the odds ratio not an appropriate estimate of the risk ratio?

Answer 5
In general, you cannot calculate a risk ratio directly from a case-control study because you do not have the denominator data from which to determine the risks for the exposed and unexposed groups. The numbers in the b and d cells represent the control group, the size of which is arbitrary. (See Example on next page.) However, if the outcome or disease of interest is rare, then cell a is small relative to b, so a/(a+b) is close to a/b; c is small relative to d, so c/(c+d) is close to c/d. The formula for the risk ratio in a cohort study can therefore be simplified as follows:

$$RR = \frac{a/a+b}{c/c+d} = (\textit{if rare disease}) = \frac{a/b}{c/d} = \frac{ad}{bc} = \textit{odds ratio}\ (OR)$$

When the outcome or disease of interest is not rare (above approximately 5%-10%), the odds ratio is not a good estimator of the risk ratio.

Instructor's Note: At CDC, the most common type of cohort study is the "church picnic"-type retrospective cohort study that has a limited exposure period. The appropriate measure of association for that type of study is the risk ratio. At CDC, the most common type of case-control study is one in which controls are selected from the population that remains disease-free. For these two types of studies, the rare disease assumption IS necessary for the odds ratio to approximate the risk ratio. However, the rare disease assumption is NOT necessary for the odds ratio to approximate the person-time type rate ratio, nor is it necessary if controls are selected at the beginning (see *case-cohort* below). Neither of these study designs are commonly used at CDC.

"It has been claimed that the odds ratio from a case-control study estimates the incidence rate ratio or the risk ratio only if the disease is rare. The reason for this belief relates to the strategy used for sampling controls. For example, in a *case-cohort study*, controls are sampled from the initial roster of all subjects, a strategy that provides a valid estimate of the risk ratio whether the disease is common or rare. **If controls were instead sampled from those who, at the end of the follow-up, remained free of disease, the odds ratio would overestimate the risk ratio** for a positive exposure-disease relation because the exposure proportion among those remaining free of disease at the end of the follow-up would be smaller than the exposure proportion among those starting their follow-up. If disease were rare, however, the odds ratio after using this sampling strategy would be a reasonable estimate of the risk ratio. In density [person-time] case-control studies or case-cohort studies, there is no need for any rare disease assumption for the odds ratio to be a valid estimate of the incidence rate ratio or the risk ratio, respectively." – Rothman KJ. *Epidemiology: an introduction*. New York: Oxford U. Press, 2002, p. 87.

EXAMPLE:
In a community of 1 million persons, 100 cases of a disease occurred. If investigators had information on each person in the community, they might have found:

		Disease	No Disease	Total	True Risk	
Entire population	Exposed	90	499,950	500,040	0.00018	OR = 9.00
	Unexposed	10	499,950	499,960	0.00002	RR = 9.00
		100	999,900	1,000,000		

The investigators could have decided to conduct a case-control study with one control per case-patient. The appropriate measure of association is the odds ratio. It is not possible to calculate risk without knowing the distribution of the exposure in the population, but the following shows what you would get if you simply divided a/a+b and c/c+d from the case-control study.

		Case	Control	Total	"Risk"	
One control per case	Exposed	90	50	140	0.643	OR = 9.00
	Unexposed	10	50	60	0.167	
	Total	100	100	200		

At four controls per case, the following would be found:

		Case	Control	Total	"Risk"	
Four controls per case	Exposed	90	200	290	0.310	OR = 9.00
	Unexposed	10	200	210	0.048	
	Total	100	400	500		

From these data it can be seen that the odds ratio is a very good approximator of the true risk ratio when the disease is rare.

The investigators used the data in Table 1 and the formulas shown below to calculate an odds ratio, a Mantel-Haenszel Chi, and 95% test-based confidence limits.

Measure	Formula	Calculation from Table 1
Odds Ratio (OR)	$OR = \frac{ad}{bc}$	$\frac{(93)(683)}{(959)(86)} = 0.77$
Expected Value of Cell 'a' (E(a))	$E(a) = \frac{H_1 V_1}{T}$	$\frac{(1052)(179)}{1821} = 103.41$
Mantel-Haenszel (MH) Variance	$MH\ variance = \frac{H_1 H_0 V_1 V_0}{T^2(T-1)}$	$\frac{(1052)(769)(179)(1642)}{(1821)^2(1821-1)} = 39.40$
Mantel-Haenszel Chi*	$X_{MH} = \frac{a - E(a)}{\sqrt{MH\ variance}}$	$\frac{93 - 103.41}{\sqrt{39.40}} = -1.66$
Test-Based Confidence Limits (Note: Z values for 2-sided confidence limits are: 90% = 1.645, 95% = 1.96, 99% = 2.58.)	$LOWER\ CL = OR^{(1-(Z/X_{MH}))}$ $UPPER\ CL = OR^{(1+(Z/X_{MH}))}$	$0.77^{(1-(1.96/-1.66))} = 0.57$ $0.77^{(1+(1.96/-1.66))} = 1.05$

* The Mantel-Haenszel Chi with one degree of freedom (X_{MH}) is equivalent to a "Z score" and may be used to find the 2-tailed p value from a table of areas in two tails of the standard normal curve. In this case, p = 0.097.

Question 6: What special information does the odds ratio give that you do not get from Chi square and p value? What additional information do you get from the p value and Chi square? From a confidence interval?

Answer 6
The OR is a measure of the direction and strength (magnitude) of the association (effect), independent of sample size.

Chi squares and p values give a feeling for how certain you are that chance does not play a role in the association--that is, the statistical significance. Both are dependent on sample size and magnitude of the association.

The confidence interval indicates the precision of the odds-ratio estimate. It can also be interpreted as the range of values with which the data are consistent. In addition, a 95% test-based confidence interval that does not include 1.0 indicates statistical significance of the association at the $p < 0.05$ level.

In summary: An odds ratio and 95% confidence interval provide a best estimate of the measure of association and the range of values with which the data are consistent. The 95% confidence intervals also permit an assessment of the role of chance. In contrast, the Chi square/p value only assess the role of chance. A non-significant Chi square/p value may indicate either that no true association exists or that the sample size is too small to detect a true association.

Question 7: How might you describe and interpret these results?

Answer 7
"These results" include:
a. the point estimate of effect or association as measured by odds ratio,
b. the precision of the point estimate as measured by the 95% confidence interval, and
c. the statistical significance of the observed association as measured by the chi and p value and/or the confidence interval's relationship to the null value of 1.0

So an interpretation should address all three.
a. Strictly speaking, the odds ratio of 0.77 indicates that the odds of a case having used OCs was about 23% less than the odds of OC use among controls. This is algebraically equivalent to the statement that the odds of an OC user developing ovarian cancer is 23% less than the odds of a non-OC user developing ovarian cancer. Since the odds ratio is a reasonable approximation of the risk ratio for a rare disease such as ovarian cancer, it is reasonable to use the term "risk" in place of "odds" in this situation. Thus one might say that the crude results are consistent with an apparent modest decreased risk of ovarian cancer associated with oral contraceptive use (OR = 0.8) of about 20%;
b. The confidence interval is relatively narrow, ranging from a 43% reduction in risk to a slight increase in risk. At the very least, it is fair to say that these results are not consistent with a substantially INCREASED risk.
c. As indicated by the confidence interval that includes the null value of 1.0 and a p value = 0.10, the association is not statistically significant at the $p<0.05$ level. So chance cannot be excluded as an explanation for the observed association, given the p value and confidence interval.

Note that, in addition, although the investigators have attempted to minimize selection bias and information bias, confounding has not been addressed.

In many epidemiologic studies, age is a confounding factor.

Question 8: What is confounding? Under what circumstances would age be a confounder in this study?

Answer 8
Confounding is the misrepresentation of an association between an exposure and a disease because of an interrelationship between the exposure, the disease, and some third factor (a "confounder").

A confounder must be:
1. associated with the exposure being studied, and
2. associated with the disease being studied independent of exposure, and
3. not a consequence of exposure, e.g., not in the causal chain.

A confounding variable is found in a particular study and is not a "fact of nature." It is undesirable in that, if it is not controlled for, it obscures the true association between exposure and disease.

So for age to be a confounder,
1. age must be associated with exposure (OC use), even among the controls.
2. age must be associated with outcome (ovarian cancer), even among the unexposed group.
3. age shouldn't be in the causal chain (not an issue here).

PART III

In the analysis of use of oral contraceptives and ovarian cancer, age was related both to OC use and to case-control status. (OC users were younger than never-users; case-patients were younger than controls.) Therefore, the investigators decided to stratify the data by age and calculate stratum-specific and, if appropriate, summary statistics of the stratified data. The Mantel-Haenszel (MH) procedure is a popular method for calculating a summary odds ratio and test of significance for stratified data.

Question 9: What is stratification? Why stratify data? How do you decide on which variables to stratify?

Answer 9
Stratification is the separation of data into subgroups of a third variable (e.g., age groups).

Stratifying data is a good method for teasing out and examining the effects of some third variable. It is used to assess and control for confounding because, by stratifying, you compare like with like (e.g., similar ages for case-patients and controls). It also allows the investigator to compare the association between exposure and outcome in different subgroups, to determine whether the effect of the exposure on the outcome varies between different subgroups -- i.e., effect modification.

Generally, one decides to stratify by variables considered to be potential confounders and effect modifiers. It is not always obvious which variables to consider as possible confounders. A reasonable approach is to identify (on the basis of literature review, subject-matter knowledge, etc.) variables that are risk factors or otherwise related to the disease in question. By definition, such variables meet one of the two criteria for a confounder. If any of the variables are also associated with the exposure in your study, then they may be confounding your results and should be controlled for. For example, age is associated with almost every disease imaginable and is associated with many exposures as well. Age, therefore, is almost always at the top of the list of potential confounders in any study.

Question 10: What is effect modification? How do you look for it?

Answer 10
Effect modification means that the degree of association between an exposure and outcome differs in different subgroups of the population. Some examples include:
1. Tetracycline and tooth mottling, modified by age.
2. Aspirin and Reye Syndrome, modified by age.
3. Heterosexual risk of human immunodeficiency virus (HIV), modified by genital ulcers.
4. Vaccine efficacy modified by age.
5. Gender and hip fracture, modified by age.
6. Measles (or diarrhea) and death, modified by HIV or nutrition.

True effect modification is not governed by the particulars of any study but is a fact of nature. Effect modification is desirable because it can help you target high-risk populations for programmatic purposes, and it may shed light on biologic mechanisms. If effect modification is present, it may be misleading to calculate a summary or "average" odds ratio, since the "average" may not apply to anyone. It is usually more meaningful to talk of the separate effects of the exposure in the different subgroups.

Effect modification is identified by calculating exposure-disease odds ratios (or risk ratios) for different subgroups and determining whether these stratum-specific odds ratios are different from one another.

Question 11a: Using the data in Table 2, calculate the odds ratio for the 40- to 49-year age stratum.

Answer 11a

Odds ratio = ad / bc = (30)(301) / (463)(30) = 0.65

Question 11b: Using the data in Table 2, calculate the expected value of cell A for the 40- to 49-year age stratum.

Answer 11b

INSTRUCTOR'S NOTE: You can ask, "What are the assumptions in calculating the expected value in a 2-by-2 table?" The assumptions are:

- the null hypothesis is correct
- the totals outside the table are fixed (can't change)

Under the null hypothesis, you would expect that the prevalence of exposure among cases and among controls would be the same, and they would be the same as the prevalence of exposure among the cases and controls combined (which is H1 / T). Assuming that the totals outside the table are fixed, then the expected prevalence of exposure among the cases would be the overall prevalence (H1/T) times the total number of cases (V1). Thus,

the expected value for cell 'a' = V1 x H1 / T = 60 x 493 / 824 = 35.90

INSTRUCTOR'S NOTE: The expected value is larger than the observed value, i.e., there are fewer exposure cases than expected. This is consistent with an odds ratio < 1, i.e., and apparently protective effect,

Question 11c: Using the data in Table 2, calculate the Mantel-Haenszel chi for the 40- to 49-year age stratum.

Answer 11c

INSTRUCTOR'S NOTE 1: Someone may ask why we are calculating an MH Chi instead of an MH Chi-square. Does it make a difference? The answer is that, no, in general, it does not. However, we are calculating a Chi instead of a Chi square to show that:

1. It is acceptable to calculate an MH Chi instead of an MH Chi square.
2. To get a p value from a Chi we would refer to a table of Z values. In general, Z tables are much more detailed than Chi square tables. A chi > 1.96 corresponds to a p value < 0.05.
3. One can see at a glance, based on whether the Chi is positive or negative, whether the odds ratio for exposure and disease is greater than or <1.0. A negative Chi reflects an odds ratio <1.0.
4. If you use the Chi in the calculation of test-based confidence limits, the lower bound will always be $OR^{(1-Z/chi)}$, and the upper bound will always be $OR^{(1+Z/chi)}$. If you use the square root of the Chi square in the formula, the lower and upper bounds will hold for odds ratios >1.0, but will be reversed for odds ratios <1.0.

INSTRUCTOR'S NOTE 2: The MH Chi, like many statistical tests, is of the form:

$$\frac{\text{observed - expected}}{\text{square root of variance}}$$

Age 40-49 years	Case	Control	Total
Ever user	30	463	493
Never user	30	301	331
Total	60	764	824

Given: cell a = 30, E(a) = 35.9, and MH variance = 13.39

Mantel-Haenszel chi = (a - E(a)) / $\sqrt{}$variance = (30 - 35.9) / $\sqrt{}$(13.39) = -1.61

Table 2. Ever-use of oral contraceptives and risk of ovarian cancer, stratified by age, Cancer and Steroid Hormone Study, 1980-1981

Ages 20-39 years

	Case	Control	Total	
Ever user	46	285	H_1 = 331	OR = 0.69
Never user	12	51	H_0 = 63	Expected(a) = 48.73
Total	V_1 = 58	V_0 = 336	T = 394	MH variance = 6.66
				MH Chi = -1.06
				95% CLs = 0.34, 1.38

Ages 40-49 years

	Case	Control	Total	
Ever user	30	463	H_1 = 493	OR = ____
Never user	30	301	H_0 = 331	Expected(a) = ____
Total	V_1 = 60	V_0 = 764	T = 824	MH variance = 13.39
				MH Chi = ____
				95% CLs = 0.38, 1.10

Ages 50-54 years

	Case	Control	Total	
Ever user	17	211	H_1 = 228	OR = 0.61
Never user	44	331	H_0 = 375	Expected(a) = 23.06
Total	V_1 = 61	V_0 = 542	T = 603	MH variance = 12.91
				MH Chi = -1.69
				95% CLs = 0.34, 1.08

The investigators had been taught to look for effect modification before looking for confounding.

Question 12: Do you think age is an effect modifier of the oral contraceptive and ovarian cancer association?

Answer 12
The odds ratios for the three strata are 0.69, 0.65, and 0.61. Since the odds ratios are very similar, age does not appear to meaningfully modify the relationship between oral contraceptives and ovarian cancer.

The investigators concluded that age was not an effect modifier. They therefore decided to control for confounding by calculating an odds ratio adjusted for age, also called a summary odds ratio or Mantel-Haenszel odds ratio, using the following formula:

$$OR_{MH} = \frac{\Sigma\ (ad/T)}{\Sigma\ (bc/T)}$$

They also calculated a Mantel-Haenszel chi, from which they found a p-value. Finally, they calculated a 95% confidence interval of 0.45 to 0.92.

Question 13a: Using the stratified data in Table 2, calculate the summary odds ratio adjusted for age.

Answer 13
Odds ratio
- numerator = (46 × 51)/394 + (30 × 301)/824 + (17 × 331)/603 = 26.2
- denominator = (285 × 12)/394 + (463 × 30)/824 + (211 × 44)/603 = 40.9
- so odds ratio = 26.2 / 40.9 = 0.64

Question 13b: Based on the Mantel-Haenszel chi of -2.55 and the attached table of the standard normal curve, determine the 2-tailed p-value.

Answer 13b
INSTRUCTOR'S NOTE: Walk the students through the use of the attached table to derive the p-value. See the first footnote at the bottom of the table.

- from table, for chi = -2.55, p-value = 0.011

Question 14: In terms of the null hypothesis and statistical significance, what do you infer from the p-value? What do you infer from the confidence interval of 0.45–0.92?

Answer 14
Compared to a cut-off ("alpha level") of 0.05, the observed p-value is clearly smaller, so one would reject the null hypothesis and conclude that there is a statistically significant association between oral contraceptives and ovarian cancer, controlling for age. The association is in the protective direction, as indicated by fewer observed than expected exposed cases, an odds ratio < 1, and a negative chi.

The 95% confidence interval, which can be interpreted as the range of values consistent with the data in this study, does not include 1.0 (which represents the null hypothesis). Therefore, as above, one would conclude that this study is not consistent with the null hypothesis.

Note that the p-value and 95% confidence interval are usually but not always consistent in terms of rejecting the null hypothesis or not. There are several different formulas for p-values and confidence intervals. As a result, occasionally, particularly when the calculated p-value is just in the neighborhood of 0.05, the p-value may be just less than 0.05 but the 95% confidence interval barely includes 1.0, or vice versa. In that situation, since the p-value is in the vicinity of 0.05, and 0.05 should not be viewed as an "on/off" switch anyway, one might say the association is "borderline" statistically significant.

Question 15: Do you think age is a confounding variable in this analysis of the association between OC use and ovarian cancer?

Answer 15
To look for confounding, compare the crude odds ratio with the adjusted odds ratio. For practical purposes, a variable may be considered a confounder when the adjusted odds ratio is "appreciably different" from the unadjusted odds ratio (often 10% 15%, or 20% is considered "appreciably different," but this is obviously subjective).

One way to assess whether confounding is present is to determine whether the crude odds ratio lies outside the range of the stratum-specific odds ratios; if so, confounding is present. In this study, the crude odds ratio (0.77) is clearly outside the range of the stratum-specific odds ratios (0.61 - 0.69), so confounding is present. It is reasonable to conclude that the crude odds ratio somewhat misrepresents the data, whereas the adjusted odds ratio (0.64) represents the data well.

A second way to assess whether confounding is present is to determine whether the crude odds ratio and the adjusted odds ratio are "appreciably different." In the CASH study, the investigators decided that age was a confounding variable because the crude OR was 0.77 and the OR_{MH} adjusted for age was 0.64, an "appreciable" change (in their judgment!) of 17%. Note also that the adjusted variance (33.0) is smaller than the crude variance (39.4), indicating that the stratified / adjusted analysis is more "statistically efficient." As a result, the adjusted confidence interval is narrower and the result has become statistically significant.

Question 16: What are the other ways of eliminating confounding in a study?

Answer 16
The five ways to eliminate confounding are:
- randomization (not applicable in observational studies)
- restriction
- matching
- stratification / summarization
- modeling

In the introduction to this case study, pregnancy was described as apparently protective against ovarian cancer. The investigators were interested in seeing whether the association between OC use and ovarian cancer differed for women of different parity. Table 3 shows parity-specific data.

Table 3. Ever-use of oral contraceptives and risk of ovarian cancer, by parity*, CASH Study, 1980-1981

Parity	Use of OCs	# Case-patients	# Controls	Age-adjusted odds ratios (95% confidence intervals)
0	Ever user	20	67	0.3 (0.1-0.8)
	Never user	25	80	
1-2	Ever user	42	369	0.8 (0.4-1.5)
	Never user	26	199	
≥3	Ever user	30	520	0.7 (0.4-1.2)
	Never user	35	400	

* Excludes seven controls (four never-users and three ever-users) and one case (ever-user) with unknown parity.

Question 17: Is there any evidence for effect modification in the data presented in Table 3?

Answer 17

Age-adjusted odds ratios in Table 3 were obtained by calculating a summary odds ratio adjusted for age, within each parity stratum. The odds ratios were age-adjusted because age was considered to be a confounder.

In Table 3, parity might be an effect modifier, since the OR = 0.3 for nulliparous women seems "appreciably different" from the ORs for parous women. This difference must be interpreted with caution because the numbers are small, and it could also be due to chance fluctuation of odds ratios among the strata.

There are statistical tests of heterogeneity to examine effect modification. The investigators performed such a test, and it was close but not quite statistically significant. However, the decision to call something an effect modifier usually comes down to judgment on the part of the investigator, in light of biologic plausibility, public health implications, etc.

AREAS IN TWO TAILS OF THE STANDARD NORMAL CURVE

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	1.000	0.992	0.984	0.976	0.968	0.060	0.952	0.944	0.936	0.928
0.1	0.920	0.912	0.904	0.897	0.889	0.881	0.873	0.865	0.857	0.849
0.2	0.841	0.834	0.826	0.818	0.810	0.803	0.795	0.787	0.779	0.772
0.3	0.764	0.757	0.749	0.741	0.734	0.726	0.719	0.711	0.704	0.697
0.4	0.689	0.682	0.674	0.667	0.660	0.653	0.646	0.638	0.631	0.624
0.5	0.617	0.610	0.603	0.596	0.589	0.582	0.575	0.569	0.562	0.555
0.6	0.549	0.542	0.535	0.529	0.522	0.516	0.509	0.503	0.497	0.490
0.7	0.484	0.478	0.472	0.465	0.459	0.453	0.447	0.441	0.435	0.430
0.8	0.424	0.418	0.412	0.407	0.401	0.395	0.390	0.384	0.379	0.373
0.9	0.368	0.363	0.358	0.352	0.347	0.342	0.337	0.332	0.327	0.322
1.0	0.317	0.312	0.308	0.303	0.298	0.294	0.289	0.285	0.280	0.276
1.1	0.271	0.267	0.263	0.258	0.254	0.250	0.246	0.242	0.238	0.234
1.2	0.230	0.226	0.222	0.219	0.215	0.211	0.208	0.204	0.201	0.197
1.3	0.194	0.190	0.187	0.184	0.180	0.177	0.174	0.171	0.168	0.165
1.4	0.162	0.159	0.156	0.153	0.150	0.147	0.144	0.142	0.139	0.136
1.5	0.134	0.131	0.129	0.126	0.124	0.121	0.119	0.116	0.114	0.112
1.6	0.110	0.107	0.105	0.103	0.101	0.099	0.097	0.095	0.093	0.091
1.7	0.089	0.087	0.085	0.084	0.082	0.080	0.078	0.077	0.075	0.073
1.8	0.072	0.070	0.069	0.067	0.066	0.064	0.063	0.061	0.060	0.059
1.9	0.057	0.056	0.055	0.054	0.052	0.051	0.050	0.049	0.048*	0.047
2.0	0.046	0.044	0.043	0.042	0.041	0.040	0.039	0.038	0.038	0.037
2.1	0.036	0.035	0.034	0.033	0.032	0.032	0.031	0.030	0.029	0.029
2.2	0.028	0.027	0.026	0.026	0.025	0.024	0.024	0.023	0.023	0.022
2.3	0.021	0.021	0.020	0.020	0.019	0.019	0.018	0.018	0.017	0.017
2.4	0.016	0.016	0.016	0.015	0.015	0.014	0.014	0.014	0.013	0.013
2.5	0.012	0.012	0.012	0.011	0.011	0.011	0.010	0.010	0.010	0.010
2.6	0.009	0.009	0.009	0.009	0.008	0.008	0.008	0.008	0.007	0.007
2.7	0.007	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.005	0.005
2.8	0.005	0.005	0.005	0.005	0.005	0.004	0.004	0.004	0.004	0.004
2.9	0.004	0.004	0.004	0.003	0.003	0.003	0.003	0.003	0.003	0.003
3.0	0.003	0.003	0.003	0.002	0.002	0.002	0.002	0.002	0.002	0.002
3.1	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.001
3.2	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
3.3	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
3.4	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.0005
3.5	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003	0.0003
3.6	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002	0.0002	0.0002
3.7	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
3.8	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
≥3.9	<0.0001									

* Use this table to find the 2-tailed p value which corresponds to a Z score or Chi (square root of chi-square) with 1 degree of freedom. For a given value of Z or chi (say, 1.98), find that value to 1 decimal place in the left-most column (1.9). The p value will be in the 1.9 row. Now find the second decimal of your Z or chi across the top row (0.08). The p value is in that column. The 2-tailed p value is at the intersection of the row and column you've identified (for 1.9 and 0.08, $p_2 = 0.048$).

To find the p value for a chi-square with 1 degree of freedom (including any chi-square from a simple 2-by-2 table, the McNemar chi-square from a matched 2-by-2 table, and the Mantel-Haenszel chi-square from stratified 2-by-2 tables), simply take the square root of the chi-square, then proceed as above.

PART IV - CONCLUSION

In their published report, the investigators wrote the following about the possible effect modification by parity:

"Parity appeared to be an effect modifier of the association between oral contraceptive use and the risk of ovarian cancer...[Table 3]. Among nulliparous women, the age-standardized odds ratio was 0.3 (95% confidence interval: 0.1-0.8). Among parous women, however, the odds ratios were closer to, but still less than, 1.0....It is possible, therefore, that oral contraceptives are most protective for women not already protected by pregnancy."

Although this case study deals with the data collected over the first 10 months (phase 1) of the study, an additional 19 months of data (phase 2) were collected and analyzed subsequently. The following table summarizes the apparent role of parity as an effect modifier in the two phases of the study.

Table 4. Age-adjusted odds ratios (aOR) and 95% confidence intervals for the association of oral contraceptive use and ovarian cancer, by parity and phase of study, CASH Study, 1980-1982

Parity	Phase 1 (months 1-10) aOR	(95% CI)	Phase 2 (months 11-29) aOR	(95% CI)	Total (months 1-29) aOR	(95% CI)
0	0.3	(0.1-0.8)	0.7	(0.5-1.2)	0.7	(0.4-1.0)
1-2	0.8	(0.4-1.5)	0.5	(0.3-0.7)	0.5	(0.4-0.8)
≥3	0.7	(0.4-1.2)	0.5	(0.4-0.8)	0.6	(0.4-0.8)
Total	0.6	(0.4-0.9)	0.5	(0.4-0.7)	0.6	(0.5-0.7)

On the basis of the full study results, it appeared that the effect of oral contraceptives on ovarian cancer was not substantially different for nulliparous women and for parous women.

Although there were no published studies of oral contraceptives and ovarian cancer when this study was launched, there were several by the time this study was published. Almost all showed an apparently protective effect of oral contraceptives on ovarian cancer.

Question 18: What are the public health and/or policy implications of the apparently protective effect of oral contraceptives on ovarian cancer?

Answer 18
Label OCs re: positive side effect? Prescribe to reduce ovarian cancer risk?

Before touting the benefits of OCs for non-contraceptive uses, one would probably want to assess both the benefits (protects against ovarian and endometrial cancer) to the risks (pulmonary embolus, etc.) of OCs.

After much debate, the Food and Drug Administration allowed OC manufacturers to modify their package inserts and note the possible beneficial side effects of OCs on ovarian cancer (and endometrial cancer). This was the first instance of FDA allowing a class of drugs to have positive side effects listed on the package inserts.

References – CASH Study

1. Centers for Disease Control. Oral contraceptive use and the risk of ovarian cancer: the Centers for Disease Control Cancer and Steroid Hormone Study. *JAMA* 1983;249:1596-9.

2. Centers for Disease Control. The reduction in risk of ovarian cancer associated with oral contraceptive use: the Cancer and Steroid Hormone Study of the Centers for Disease Control and the National Institute of Child Health and Human Development. *N Engl J Med* 1987;316:650-5.

References - Data Analysis, Confounding, Effect Modification

3. Dicker RC. Analyzing and Interpreting Data. In: Gregg MB. *Field Epidemiology, 2nd edition*. New York: Oxford University Press, 2002.

4. Rothman KJ. *Epidemiology: an introduction.* New York: Oxford University Press, 2002.